A.1 Fonctions gamma et bêta

A.1.1 Fonction gamma

1. Pour tout $x > 0$, la fonction gamma $\Gamma(x)$ est définie par :

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t}\, dt\,. \tag{2A.1}$$

2. En intégrant par partie, on montre que :

$$\Gamma(x+1) = x\,\Gamma(x)\,. \tag{2A.2}$$

3. En applicant cette formule de récurrence à des entiers $x = n$:

$$\Gamma(n) = (n-1)!\,. \tag{2A.3}$$

4. Il sera parfois utile d'utiliser la valeur particulière :

$$\Gamma(1/2) = \sqrt{\pi} \tag{2A.4}$$

d'où on déduit, par exemple, et en utilisant (2A.2) : $\Gamma(3/2) = (1/2)\sqrt{\pi}$.

A.1.2 Quelques intégrales utiles liées à la fonction gamma

Pour tout $a > 0$ et $p > 0$, on a :

$$\int_0^{\infty} x^{p-1} e^{-ax}\, dx = a^{-p}\,\Gamma(p) \tag{2A.5}$$

$$\int_0^{\infty} x^{-(p+1)} e^{-ax^{-1}}\, dx = a^{-p}\,\Gamma(p)\,. \tag{2A.6}$$

Les deux intégrales qui suivent pourront aussi être utiles ; elles s'obtiennent par simple transformation de variables :

$$\int_0^{\infty} x^{p-1} e^{-ax^2}\, dx = \frac{1}{2} a^{-p/2}\,\Gamma\left(\frac{p}{2}\right) \tag{2A.7}$$

$$\int_0^{\infty} x^{-(p+1)} e^{-ax^{-2}}\, dx = \frac{1}{2} a^{-p/2}\,\Gamma\left(\frac{p}{2}\right). \tag{2A.8}$$

A.1.3 Fonction bêta et intégrale de Dirichlet

1. Pour tout $a, b > 0$, la fonction bêta est définie par l'intégrale :

$$B(a,b) = \int_0^{1} t^{a-1}(1-t)^{b-1}\, dt\,. \tag{2A.9}$$

2. Son lien avec la fonction gamma est donné par :

$$B(a,b) = \frac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)}\,. \tag{2A.10}$$

3. L'intégrale définissant la fonction bêta en (2A.9) se généralise au cas multivarié : c'est l'intégrale de Dirichlet :

$$\int_{S_k} t_1^{a_1-1} \dots t_k^{a_k-1} (1 - t_1 - \dots - t_k)^{a_{k+1}-1} \, dt_1 \dots dt_k = \frac{\Gamma(a_1) \dots \Gamma(a_{k+1})}{\Gamma(a_1 + \dots + a_{k+1})} \tag{2A.11}$$

où $a_j > 0, \ j = 1, \dots, k+1$ et S_k est un simplexe dans $\mathbb{R}^k$:

$$S_k = \left\{ (t_1, \dots, t_k) \in \mathbb{R}^k \mid t_j \geqslant 0, \ j = 1, \dots, k, \ \sum_{j=1}^{k} t_j \leqslant 1 \right\}. \tag{2A.12}$$

Dans le cas particulier $k = 1$, on retrouve l'intégrale bêta.

A.2 Lois univariées de base

A.2.1 La loi discrète binomiale

La variable aléatoire discrète binomiale de paramètres (n, π), $n \geqslant 1$, $0 \leqslant \pi \leqslant 1$ est caractérisée par sa fonction de probabilité :

$$p(x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x} \mathbf{1}_{0,1,\dots,n}(x) \,. \tag{2A.13}$$

On notera aussi cette fonction $f_b(x \mid n, \pi)$ et on écrira plus simplement :

$$x \sim \mathrm{Bin}\,(n, \pi)\,.$$

Les moments sont donnés par :

$$\begin{cases} \mathrm{E}\,(x) &= n\pi \\ \mathrm{Var}(x) &= n\pi\,(1 - \pi)\,. \end{cases} \tag{2A.14}$$

La loi binomiale apparaît lorsqu'on s'intéresse au nombre de succès dans une suite de n épreuves de Bernoulli (succès/échecs) indépendantes où π est la probabilité de succès à chaque épreuve.

Lorsque $n = 1$, la loi (2A.13) est d'ailleurs appelée la loi de Bernoulli de paramètre π et on écrira $x \sim \mathrm{Bern}(\pi)$. On a :

$$p(x) = \pi^x (1 - \pi)^{1-x} \mathbf{1}_{0,1}(x)\,. \tag{2A.15}$$

A.2.2 La loi discrète de Poisson

La variable aléatoire discrète de Poisson de paramètres $\lambda > 0$ est caractérisée par sa fonction de probabilité :

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!} \mathbf{1}_{0,1,2,\dots}(x)\,. \tag{2A.16}$$

On notera aussi cette fonction $f_{Poi}(x \mid \lambda)$ et on écrira :

$$x \sim \mathrm{Poisson}(\lambda)\,.$$

Les moments sont donnés par :

$$\begin{cases} \mathrm{E}(x) &= \lambda \\ \mathrm{Var}(x) &= \lambda\,. \end{cases} \tag{2A.17}$$

Cette loi sera utile dans des processus de comptage où, sous certaines hypothèses, le nombre de fois x qu'un événement se produit par unité de temps, de longueur, de surface ou de volume, a une distribution de Poisson.

A.2.3 Densité normale univariée

La densité de x est donnée par :

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \tag{2A.18}$$

où $\mu \in \mathbb{R}$ et $\sigma > 0$. Cette densité sera notée $f_N(x \mid \mu, \sigma^2)$ et on écrira plus simplement :

$$x \sim N(\mu, \sigma^2).$$

On a évidemment pour les premiers moments :

$$\begin{cases} \mathrm{E}(x) &= \mu \\ \mathrm{Var}(x) &= \sigma^2\,. \end{cases} \tag{2A.19}$$

On se rappelera aussi qu'une transformation linéaire d'une variable normale définit une nouvelle variable normale : si $y = ax + b$, alors $y \sim N(a\mu + b, a^2\sigma^2)$. En particulier, cette propriété permet de calculer des probabilités sur des normales générales $N(\mu, \sigma^2)$ à partir de tables de la normale standardisée $N(0, 1)$.

A.2.4 Les densités gamma et associées

A.2.4.1 Densité gamma

La densité d'une variable gamma x est donnée par :

$$p(x) = \frac{\lambda^\alpha x^{\alpha-1} \exp\{-\lambda x\}}{\Gamma(\alpha)} \tag{2A.20}$$

où $x \geqslant 0$ et $\lambda, \alpha > 0$. On utilisera aussi la notation $p(x) = f_\gamma(x \mid \lambda, \alpha)$ ou encore :

$$x \sim \Gamma(\lambda, \alpha)\,.$$

On a pour les premiers moments :

$$\begin{cases} \mathrm{E}(x) &= \dfrac{\alpha}{\lambda} \\ \mathrm{Var}(x) &= \dfrac{\alpha}{\lambda^2}\,. \end{cases} \tag{2A.21}$$

A partir de cette densité, une série d'autres densités utiles ont été définies, soit comme cas particuliers, soit comme transformations simples.

A.2.4.2 Densité exponentielle

C'est un cas particulier de la densité gamma, où $\alpha = 1$. La densité de x est donc donnée par :

$$p(x) = \lambda \exp\{-\lambda x\} \tag{2A.22}$$

où $x \geqslant 0$ et $\lambda > 0$. On écrira aussi :

$$x \sim \mathrm{Exp}\,\{\lambda\}\,.$$

Les moments de x sont :

$$\left\{\begin{array}{rcl} \mathrm{E}(x) & = & \dfrac{1}{\lambda} \\ \mathrm{Var}(x) & = & \dfrac{1}{\lambda^2}\,. \end{array}\right. \tag{2A.23}$$

A.2.4.3 Densité khi-carré

La khi-carré à ν degrés de liberté est un cas particulier de la gamma : $\chi^2_\nu \sim \Gamma(1/2, \nu/2)$. La densité est donnée par :

$$p(x) = \frac{(1/2)^{\nu/2} x^{\nu/2-1} \exp\{-(1/2)x\}}{\Gamma\,(\nu/2)} \tag{2A.24}$$

où $x \geqslant 0$ et $\nu > 0$. On utilisera la notation $p(x) = f_{\chi^2}(x \mid \nu)$ ou encore :

$$x \sim \chi^2_\nu\,.$$

Les deux premiers moments de x sont les suivants :

$$\left\{\begin{array}{rcl} \mathrm{E}(x) & = & \nu \\ \mathrm{Var}(x) & = & 2\nu\,. \end{array}\right. \tag{2A.25}$$

On peut aussi vérifier la propriété suivante :

$$\text{Si } x \sim \Gamma\left(\frac{\lambda}{2}, \frac{\nu}{2}\right) \text{ alors } y = \lambda\, x \sim \chi^2_\nu\,. \tag{2A.26}$$

A.2.4.4 Densité gamma-inverse

Une gamma-inverse est la distribution de probablilté de l'inverse d'une variable aléatoire gamma. Soit $y \sim \Gamma(\lambda, \nu)$ et $x = 1/y$. La densité de x est donnée par :

$$p(x) = \frac{\lambda^\alpha (1/x)^{\alpha+1} \exp\{-\lambda/x\}}{\Gamma(\alpha)} \tag{2A.27}$$

où $x \geqslant 0$ et $\lambda, \alpha > 0$. On écrira aussi $p(x) = f_{i\gamma}(x \mid \lambda, \alpha)$ ou encore :

$$x \sim \Gamma^{-1}\,(\lambda, \alpha)\,.$$

On a pour les premiers moments :

$$\left\{\begin{array}{rcll} \mathrm{E}(x) & = & \dfrac{\lambda}{\alpha - 1} & (\alpha > 1) \\ \mathrm{Var}(x) & = & \dfrac{\lambda^2}{(\alpha-1)^2(\alpha-2)} & (\alpha > 2)\,. \end{array}\right. \tag{2A.28}$$

A.2.4.5 Densité khi-carré-inverse

De même, une khi-carré-inverse est la loi de l'inverse d'une variable khi-carré. C'est donc un cas particulier de la gamma-inverse $\chi_\nu^{-2} \sim \Gamma^{-1}(1/2, \nu/2)$. La densité de x est donnée par :

$$p(x) = \frac{(1/2)^{\nu/2}(1/x)^{\nu/2+1}\exp\{-1/2x\}}{\Gamma(\nu/2)} \tag{2A.29}$$

où $x \geqslant 0$ et $\nu > 0$. On utilisera ici aussi la notation $p(x) = f_{\chi^{-2}}(x \mid \nu)$ et :

$$x \sim \chi_\nu^{-2}.$$

On a pour les premiers moments :

$$\left\{\begin{array}{lcll} \mathrm{E}(x) & = & \dfrac{1}{\nu-2} & (\nu > 2) \\ \mathrm{Var}(x) & = & \dfrac{2}{(\nu-2)^2(\nu-4)} & (\nu > 4)\,. \end{array}\right. \tag{2A.30}$$

On a ici la propriété :

$$\text{Si } x \sim \Gamma^{-1}\left(\frac{\lambda}{2}, \frac{\nu}{2}\right) \text{ alors } y = \frac{x}{\lambda} \sim \chi_\nu^{-2}\,. \tag{2A.31}$$

A.2.5 Densité de Student généralisée univariée

La loi de Student généralisée apparaît lorsqu'on mélange des lois normales $x \sim N(\mu, \sigma^2)$ pour différentes valeurs de σ^2 distribuées selon une loi gamma-inverse « mélangeante » soit $\sigma^2 \sim \Gamma^{-1}[(\beta\nu/2), (\nu/2)]$. La loi résultante de x va donc dépendre de trois paramètres : (μ, β, ν). On obtient :

$$\begin{aligned} p(x) &= \int_0^\infty f_N\left(x \mid \mu, \sigma^2\right) f_{i\gamma}\left(\sigma^2 \,\middle|\, \frac{\beta\nu}{2}, \frac{\nu}{2}\right) d\sigma^2 \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \\ &\quad \times \frac{(\beta\nu/2)^{\nu/2}(1/\sigma^2)^{\nu/2+1}\exp\left\{-\beta\nu/2\sigma^2\right\}}{\Gamma(\nu/2)}\, d\sigma^2\,. \end{aligned} \tag{2A.32}$$

En utilisant (2A.4) et (2A.6), on obtient le résultat :

$$p(x) = \frac{1}{B\left[(1/2), (\nu/2)\right]}\,(\beta\nu)^{-1/2}\left(\frac{(x-\mu)^2}{\beta\nu}+1\right)^{[-(1/2)(\nu+1)]} \tag{2A.33}$$

où $x, \mu \in \mathbb{R}$ et $\beta, \nu > 0$. Cette densité est appelée la densité Student généralisée de paramètres (μ, β, ν). On écrira souvent $p(x) = f_t(x \mid \mu, \beta, \nu)$ ou encore :

$$x \sim t\,(\mu, \beta, \nu)\,.$$

On trouve parfois cette densité écrite sous la forme équivalente :

$$p(x) = \frac{1}{B\,[(1/2), (\nu/2)]}\,\nu^{\nu/2}\beta^{-1/2}\left(\frac{(x-\mu)^2}{\beta}+\nu\right)^{[-(1/2)(\nu+1)]} \tag{2A.34}$$

On a pour les premiers moments :

$$\begin{cases} \mathrm{E}(x) &= \mu \quad (\nu > 1) \\ \mathrm{Var}(x) &= \beta\dfrac{\nu}{\nu-2} \quad (\nu > 2)\,. \end{cases} \tag{2A.35}$$

On verra plus loin que cette loi univariée a son équivalent multivarié qui joue un rôle important dans l'analyse des modèles de régression.

Cas particulier : la Student-t à ν degrés de liberté

Si on choisit $\mu = 0$ et $\beta = 1$, on retrouve la loi de Student-t à ν degrés de liberté $t(0, 1, \nu) = t_\nu$. En fait, si $x \sim t(\mu, \beta, \nu)$, alors :

$$y = \frac{x-\mu}{\sqrt{\beta}} \sim t_\nu\,.$$

On a :

$$p(y) = \frac{1}{B\,[(1/2), (\nu/2)]}\,\nu^{-1/2}\left(\frac{y^2}{\nu}+1\right)^{[-(1/2)(\nu+1)]}\,. \tag{2A.36}$$

Cette relation permet d'utiliser les tables de Student-t standardisée $t_\nu = t(0, 1, \nu)$, que l'on trouve dans tous les manuels de base de statistique, pour calculer des probabilités relatives à une Student généralisée $t(\mu, \beta, \nu)$. On utilisera des transformations similaires à celles qui permettent de passer d'une $N(\mu, \sigma^2)$ à une $N(0, 1)$.

A.2.6 Densité bêta

La densité bêta est définie pour une variable aléatoire continue sur l'intervalle $[0, 1]$. Elle sera donc particulièrement utile pour modéliser l'incertitude sur toute variable prennant ses valeurs dans l'intervalle $[0, 1]$ comme, par exemple, une proportion. La densité bêta est définie par :

$$p(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\,x^{a-1}(1-x)^{b-1}\mathbf{1}_{[0,1]}(x)\,, \tag{2A.37}$$

où $a, b > 0$. On notera cette densité $f_\beta(x \mid a, b)$ et on écrira plus simplement :

$$x \sim \text{Beta}\,(a, b)\,.$$

Les premiers moments sont :

$$\begin{cases} \mathrm{E}(x) &= \dfrac{a}{a+b} \\ \mathrm{Var}(x) &= \dfrac{ab}{(a+b)^2(a+b+1)}\,. \end{cases} \tag{2A.38}$$

Un cas particulier intéressant est l'« uniforme continue » sur $[0, 1]$. Il suffit de prendre $a = b = 1$.

A.2.7 La loi discrète négative-binomiale

La loi négative-binomiale apparaît dans une suite d'épreuves de Bernoulli (succès / échecs) indépendantes où π est la probabilité de succès à chaque épreuve, lorsqu'on s'intéresse au nombre d'échecs avant d'obtenir le r-ième succès(13).

La fonction de probabilité s'écrit :

$$p(x) = \binom{x+r-1}{r-1} \pi^r (1-\pi)^x \, \mathbf{1}_{\{0,1,2,\dots\}}(x) \tag{2A.39}$$

où $r \geqslant 1$ et $0 \leqslant \pi \leqslant 1$. On notera pour cette fonction $f_{nb}(x \mid r, \pi)$ et on écrira aussi :

$$x \sim \text{Negbin}(r, \pi).$$

On a pour les moments :

$$\left\{ \begin{array}{lcl} \text{E}(x) & = & \dfrac{r(1-\pi)}{\pi} \\ \\ \text{Var}(x) & = & \dfrac{r(1-\pi)}{\pi^2}. \end{array} \right. \tag{2A.40}$$

Cas particulier : si $r = 1$ on retrouve la loi géométrique de paramètre π.

Remarquons que cette loi apparaît également lorsqu'on mélange une loi de Poisson pour x par une loi gamma pour le paramètre λ de la Poisson. On peut en effet montrer, en utilisant l'intégrale gamma, que :

$$f_{nb}(x \mid r, \pi = \frac{s}{s+1}) = \int_0^\infty f_{Poi}(x \mid \lambda) \, f_\gamma(\lambda \mid s, r) \, d\lambda.$$

Ceci montre que la fonction de probabilité négative-binomiale est également parfaitement définie pour des valeurs non entières et positives de r même si son interprétation naturelle suppose que r est entier et $r \geqslant 1$. On utilisera donc parfois la loi négative-binomiale (2A.39) sous sa forme plus générale où simplement $r > 0$.

A.2.8 La loi discrète bêta-binomiale

Cette densité apparaît lorsqu'on envisage pour la variable aléatoire x un mélange de lois binomiales $\text{Bin}(n, \pi)$ pour différentes valeurs de π distribuées selon une loi $\text{Beta}(a, b)$:

$$f_{\beta b}(x \mid n, a, b) = \int_0^1 f_b(x \mid n, \pi) f_\beta(\pi \mid a, b) \, d\pi \,. \tag{2A.41}$$

La fonction de probabilité est obtenue en résolvant cette intégrale. En utilisant les propriétés de l'intégrale bêta (2A.9) on obtient :

$$p(x) = \frac{\Gamma(x+a)\,\Gamma(n-x+b)\,\Gamma(n+1)\,\Gamma(a+b)}{\Gamma(x+1)\,\Gamma(n-x+1)\,\Gamma(a)\,\Gamma(b)\,\Gamma(n+a+b)} \tag{2A.42}$$

où $x = 0, 1, 2, \dots, n$; les paramètres sont $n = 1, 2, \dots$ et $a, b > 0$. On écrira plus simplement :

$$x \sim \text{BetaBin}\,(n, a, b)\,.$$

(13) La loi discrète de Pascal donne les probabilités pour le nombre d'*essais nécessaires* pour obtenir le r-ième succès. Si $x \sim \text{Negbin}(r, \pi)$ alors, $y = x + r$ suit une loi Pascal de paramètres (r, π).

Les moments sont donnés par :

$$\begin{cases} \mathrm{E}(x) &= n \dfrac{a}{a+b} \\ \mathrm{Var}(x) &= n\,(n+a+b)\, \dfrac{ab}{(a+b)^2(a+b+1)} \,. \end{cases} \tag{2A.43}$$

On remarquera que le cas particulier $a = b = 1$ assigne pour x une probabilité uniforme discrète sur les entiers $0, 1, \ldots, n$. On a en effet :

$$f_{\beta b}(x \mid n, 1, 1) = \frac{1}{n+1} \mathbf{1}_{0,1,\ldots,n}(x)\,. \tag{2A.44}$$

A.2.9 Densité bêta-inverse-1

Cette densité apparaît lorsqu'on transforme une variable $y \sim \mathrm{Beta}(a, b)$ en $x = c/y$ où $c \geqslant 0$. La densité de x s'écrit :

$$p(x) = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} \frac{c^a (x-c)^{b-1}}{x^{a+b}} \tag{2A.45}$$

où $a, b > 0$, $c \geqslant 0$ et $x \geqslant c$. On la notera $f_{i\beta 1}(x \mid a, b, c)$ et on écrira souvent :

$$x \sim \mathrm{Betainv\,1}\,(a, b, c)\,.$$

Les moments sont donnés par :

$$\begin{cases} \mathrm{E}(x) &= c\, \dfrac{a+b-1}{a-1} \quad (a > 1) \\ \mathrm{Var}(x) &= c^2\, \dfrac{(a+b-1)\,b}{(a-1)^2(a-2)} \quad (a > 2)\,. \end{cases} \tag{2A.46}$$

A.2.10 Densité bêta-inverse-2

Cette densité apparaît aussi lorsqu'on transforme une variable $y \sim \mathrm{Beta}(a, b)$, mais ici en $x = c\{y/(1-y)\}$ où $c > 0$. La densité de x s'écrit :

$$p(x) = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} \frac{c^b x^{a-1}}{(x+c)^{a+b}} \tag{2A.47}$$

où $a, b, c > 0$, et $x \geqslant 0$. On la notera $f_{i\beta 2}(x \mid a, b, c)$ pour la fonction de densité et on écrira :

$$x \sim \mathrm{Betainv\,2}\,(a, b, c)\,.$$

Les moments sont donnés par :

$$\begin{cases} \mathrm{E}(x) &= c\, \dfrac{a}{b-1} \quad (b > 1) \\ \mathrm{Var}(x) &= c^2\, \dfrac{a(a+b-1)}{(b-1)^2(b-2)} \quad (b > 2)\,. \end{cases} \tag{2A.48}$$

On remarquera que cette loi s'obtient aussi en mélangeant une loi gamma par une autre loi gamma de la façon suivante :

$$f_{i\beta 2}(x \mid a, b, c) = \int_0^\infty f_\gamma(x \mid z, a)\, f_\gamma(z \mid c, b)\, dz\,.$$

Cette propriété sera particulièrement très utile dans l'analyse bayésienne des processus gamma.

Cas particulier, la densité de Fisher-Snedecor

Si l'on prend une densité bêta-inverse-2 particulière, où $a = \nu_1/2$, $b = \nu_2/2$, et $c = \nu_2/\nu_1$, on retrouve la densité de Fisher-Snedecor à ν_1 et ν_2 degrés de liberté notée F_{ν_1,ν_2}.

A.3 Lois multinomiale et Dirichlet

A.3.1 La loi discrète multinomiale

La loi discrète multinomiale de paramètres $(n, \pi_1, \ldots, \pi_k)$ est caractérisée par la fonction de probabilité k-variée :

$$\begin{aligned} p(n_1, \ldots, n_k) &= \frac{n!}{n_1! \ldots n_k!(n - n_1 - \ldots - n_k)!} \\ &\quad \times \pi_1^{n_1} \ldots \pi_k^{n_k} (1 - \pi_1 - \ldots - \pi_k)^{n-n_1-\ldots-n_k}. \end{aligned} \tag{2A.49}$$

Les paramètres $\pi_1, \ldots, \pi_k$ appartiennent au simplexe $S_k \subset \mathbb{R}^k$:

$$S_k = \left\{ (\pi_1, \ldots, \pi_k) \in \mathbb{R}^k \mid \pi_j \geqslant 0,\ j = 1, \ldots, k\ ;\ \sum_{j=1}^{k} \pi_j \leqslant 1 \right\}.$$

Cette loi apparaît dans une suite de n observations indépendantes d'une expérience, où l'on retient le nombre n_j de fois qu'un événement A_j s'est réalisé, $j = 1, \ldots, k+1$ où $(A_1, \ldots, A_{k+1})$ forme une partition de l'ensemble des résultats possibles de l'expérience et $\pi_j = \text{Prob}(A_j)$ $j = 1, \ldots, k+1$. On a donc :

$$\begin{aligned} n_j &= \#\left(A_j\right) \quad j = 1, \ldots, k+1 \\ n_j &\geqslant 0, \quad \sum_{j=1}^{k+1} n_j = n. \end{aligned} \tag{2A.50}$$

On écrira simplement pour représenter un vecteur multinomial de densité (2A.49) :

$$n_1, \ldots, n_k \sim \text{Mult}\,(n, \pi_1, \ldots, \pi_k).$$

On a les moments suivants pour $i \neq j = 1, \ldots, k$:

$$\left\{ \begin{aligned} \text{E}(n_i) &= n\pi_i \\ \text{Var}(n_i) &= n\pi_i(1 - \pi_i) \\ \text{Cov}(n_i, n_j) &= -n\pi_i\pi_j. \end{aligned} \right. \tag{2A.51}$$

On remarquera que, dans le cas particulier $k = 1$, on retrouve la loi binomiale donnée en (2A.13), c'est-à-dire : $f_b(n_1 \mid n, \pi_1)$. On peut également montrer que les marginales de n_i sont des binomiales :

$$n_i \sim \text{Bin}(n, \pi_i) \quad (i = 1, \ldots, k).$$

3.2 Densité Dirichlet

La densité Dirichlet est une généralisation multivariée de la densité bêta. Ici, x est continue multivariée $x = (x_1, \ldots, x_k)$; elle appartient au simplexe de $\mathbb{R}^k$. La densité s'écrit :

$$p(x_1, \ldots, x_k) \tag{2A.52}$$
$$= \frac{\Gamma(a_1 + \ldots + a_{k+1})}{\Gamma(a_1) \ldots \Gamma(a_{k+1})} x_1^{a_1 - 1} \ldots x_k^{a_k - 1} (1 - x_1 - \ldots - x_k)^{a_{k+1} - 1} \mathbf{1}_{S_k}(x_1, \ldots, x_k) .$$

où :

$$S_k = \left\{ (x_1, \ldots, x_k) \in \mathbb{R}^k \mid x_i \geqslant 0,\ i = 1, \ldots, k,\ \sum_{i=1}^{k} x_i \leqslant 1 \right\}$$

est le simplexe de dimension k et ses paramètres sont $\{a_1, \ldots, a_{k+1} \mid a_j > 0,\ j = 1, \ldots, k+1\}$. On écrira parfois pour cette densité $f_{Dir}(x_1, \ldots, x_k \mid a_1, \ldots, a_{k+1})$ ou plus simplement :

$$x \sim \mathrm{Dir}\,(a_1, \ldots, a_{k+1}) .$$

Les moments principaux sont, pour $i \neq j = 1, \ldots, k$:

$$\left\{ \begin{array}{rcl} \mathrm{E}(x_i) & = & \dfrac{a_i}{a_1 + \ldots + a_{k+1}} \\[2ex] \mathrm{Var}(x_i) & = & \dfrac{a_i(a_1 + \ldots + a_{k+1} - a_i)}{(a_1 + \ldots + a_{k+1})^2 (a_1 + \ldots + a_{k+1} + 1)} \\[2ex] \mathrm{Cov}(x_i, x_j) & = & -\dfrac{a_i a_j}{(a_1 + \ldots + a_{k+1})^2 (a_1 + \ldots + a_{k+1} + 1)} . \end{array} \right. \tag{2A.53}$$

Si $k = 1$, on retrouve en effet la densité $\mathrm{Beta}(a_1, a_2)$.

A.4 Densités normale multivariée et associées

La distribution normale multivariée joue un rôle fondamental en analyse statistique, elle est décrite dans tous les ouvrages de base de la statistique multivariée (voir par exemple Morrison,1990). En analyse bayésienne elle est associée à d'autres lois très utiles : la normale-gamma multivariée et la Student multivariée.

A.4.1 Densité normale multivariée

Soit un vecteur aléatoire $x \in \mathbb{R}^m$ de moyenne $\mu \in \mathbb{R}^m$ et de matrice de covariances $\Sigma : (m \times m) > 0$. La variable x est une normale multivariée de paramètres (μ, Σ) lorsque sa densité de probabilité peut s'écrire :

$$p(x) = \frac{1}{(2\pi)^{m/2} \mid \Sigma \mid^{1/2}} \exp \left\{ -\frac{1}{2} (x - \mu)' \Sigma^{-1} (x - \mu) \right\} , \tag{2A.54}$$

où $\mid \Sigma \mid$ est le déterminant de Σ. On écrira plus simplement :

$$x \sim N_m\,(\mu,\ \Sigma) .$$

La densité normale univariée $N(\mu, \sigma^2)$ correspond au cas $m = 1$.

A.4.2 Transformations linéaires de normales

La loi normale est « stable » aux transformations linéaires. Soit une variable aléatoire $x \sim N_m(\mu,\ \Sigma)$ et une transformation linéaire de x, $y = Ax + b$ où $A : (p \times m)$ et $b : (p \times 1)$, alors :

$$y \sim N_p \left(A\mu + b,\ A\Sigma A'\right). \tag{2A.55}$$

Pour éviter de manipuler des lois normales singulières, on suppose ici que $p \leqslant m$.

A.4.3 Marginales et conditionnelles des variables normales

Soit $x = (x_1', x_2')' \in \mathbb{R}^m$ où $x_i \in \mathbb{R}^{p_i}$, $i = 1, 2$ et $p_1 + p_2 = m$ où $x \sim N_m(\mu,\ \Sigma)$. Si on partitionne μ et Σ comme le vecteur x :

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix},\ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \tag{2A.56}$$

on obtient :

$$\left\{ \begin{array}{rcl} x_1 & \sim & N_{p_1}(\mu_1,\ \Sigma_{11}) \\ x_2 \mid x_1 & \sim & N_{p_2}(\mu_{2\cdot 1},\ \Sigma_{22\cdot 1}) \end{array} \right. \tag{2A.57}$$

où :

$$\mu_{2\cdot 1} = \mu_2 + \Sigma_{21}(\Sigma_{11})^{-1}(x_1 - \mu_1) \tag{2A.58}$$

$$\Sigma_{22\cdot 1} = \Sigma_{22} - \Sigma_{21}(\Sigma_{11})^{-1}\Sigma_{12}\,. \tag{2A.59}$$

A.4.4 Recompositions de normales conjointes

Un autre résultat très utile est la recomposition d'une conjointe à partir de marginales et de conditionnelles. Supposons que l'on ait les décompositions suivantes :

$$\left\{ \begin{array}{rcl} x_1 & \sim & N_{p_1}(\mu_1,\ \Sigma_{11}) \\ x_2 \mid x_1 & \sim & N_{p_2}\,(A\,x_1 + b,\ \Omega)\,, \end{array} \right. \tag{2A.60}$$

alors on a conjointement une normale $(p_1 + p_2)$-variée :

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N_{p_1+p_2}\left[\begin{pmatrix} \mu_1 \\ A\,\mu_1 + b \end{pmatrix},\ \begin{pmatrix} \Sigma_{11} & \Sigma_{11}\,A' \\ A\,\Sigma_{11} & A\,\Sigma_{11}\,A' + \Omega \end{pmatrix}\right]. \tag{2A.61}$$

Pour que cette recomposition fonctionne, il est important de noter que :

1. la variance conditionnelle de $x_2 \mid x_1$ doit être constante : Ω (homoscédasticité) ;
2. la moyenne conditionnelle de $x_2 \mid x_1$ doit être linéaire en x_1 : $A\,x_1 + b$ (régression linéaire).

A.4.5 Echantillonnage et vraisemblance pour une normale multivariée

Soit un échantillon aléatoire i.i.d. de variables $N_m(\mu,\ \Sigma)$. La matrice des données $\mathbf{X} : (n \times m)$ s'écrit :

$$\mathbf{X} = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \vdots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}. \tag{2A.62}$$

Les estimateurs sans biais usuels des paramètres (μ, Σ) sont donnés par $(\bar{x}, S)$, où :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{2A.63}$$

$$S = \frac{A}{n-1}, \tag{2A.64}$$

où la matrice A des sommes de carrés et des produits croisés est donnée par :

$$A = \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})' \tag{2A.65}$$

$$= \sum_{i=1}^{n} x_i x_i' - n \bar{x} \bar{x}' . \tag{2A.66}$$

La fonction de vraisemblance s'écrit :

$$L(\mu, \Sigma) = \frac{1}{(2\pi)^{nm/2} \, |\Sigma|^{n/2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^{n} (x_i - \mu)' \Sigma^{-1} (x_i - \mu) \right\} \tag{2A.67}$$

$$= \frac{1}{(2\pi)^{nm/2} \, |\Sigma|^{n/2}} \exp \left\{ -\frac{1}{2} \left(n(\bar{x} - \mu)' \Sigma^{-1} (\bar{x} - \mu) + \operatorname{tr} A \, \Sigma^{-1} \right) \right\} .$$

2A.6 Densité normale-gamma multivariée

Soit un vecteur aléatoire $x \in \mathbb{R}^p$, $p \geqslant 1$ et une variable aléatoire $y > 0$. La normale-gamma se construit en combinant la marginale de y (une densité gamma particulière) et la conditionnelle de $x \mid y$ (une densité normale p-variée particulière). Si on a :

$$x \mid y \sim N_p(\mu, y^{-1} B) , \quad \mu \in \mathbb{R}^p, \; B : (p \times p) , B > 0$$

$$y \sim \Gamma\left(\frac{v \nu}{2}, \frac{\nu}{2} \right), \quad v, \nu > 0 ,$$

on écrira alors :

$$(x, y) \sim N_p \gamma \, (\mu, B, v, \nu) .$$

On a les moments suivants :

$$\begin{cases} \mathrm{E}(x \mid y) = \mu \\ \mathrm{Var}(x \mid y) = y^{-1} B \end{cases} \tag{2A.68}$$

$$\begin{cases} \mathrm{E}(y) = \dfrac{1}{v} \\ \mathrm{Var}(y) = \dfrac{2}{\nu v^2} \cdot \end{cases} \tag{2A.69}$$

On voit que B^{-1} peut s'interpréter comme une précision relative de x en unités y.

On peut calculer les moments marginaux du vecteur x :

$$\begin{cases} \mathrm{E}(x) &= E_y\big(E\,(x \mid y)\big) = \mu \\ \mathrm{Var}(x) &= E_y\big(V\,(x \mid y)\big) + V_y\big(E\,(x \mid y)\big) = E_y\left(y^{-1}\right)\, B = \dfrac{\nu}{\nu - 2}\, \upsilon B \end{cases} \tag{2A.70}$$

où $\nu > 2$ pour que la variance existe. La dernière égalité vient du fait que,

$$\text{si } y \sim \Gamma\left(\frac{\upsilon\nu}{2}, \frac{\nu}{2}\right), \text{ alors } y^{-1} \sim \Gamma^{-1}\left(\frac{\upsilon\nu}{2}, \frac{\nu}{2}\right).$$

La densité normale-gamma multivariée est donc :

$$\begin{aligned} p(x, y) &= f_{N\gamma}\,(x, y \mid \mu, B, \upsilon, \nu) \\ &= f_N\left(x \mid \mu, y^{-1}B\right)\, f_\gamma\left(y \,\Big|\, \frac{\upsilon\nu}{2}, \frac{\nu}{2}\right) \\ &= \frac{1}{(2\pi)^{r/2}}\, |B|^{-1/2} y^{r/2} \exp\left\{-\frac{1}{2} y (x - \mu)' B^{-1} (x - \mu)\right\} \\ &\quad \times \frac{(\upsilon\nu/2)^{\nu/2}\, y^{\nu/2 - 1}\, e^{-y\upsilon\nu/2}}{\Gamma(\nu/2)}\, \mathbf{1}_{\mathbb{R}^p}(x)\, \mathbf{1}_{\mathbb{R}^+}(y)\,, \end{aligned} \tag{2A.71}$$

où r est une notation pour représenter le rang de la matrice B. En fait, ce rang doit être égal à p pour que la densité normale conditionnelle de $x \mid y$ existe (il faut que B soit régulière). Ce sera en général le cas.

Remarque A1 **(Normale-Gamma impropre)**

Il sera parfois utile de considérer en analyse bayésienne des cas limites où la loi de (x, y) tend vers la loi impropre (non intégrable) :

$$p(x, y) \propto \frac{1}{y}\,.$$

On obtient ainsi pour $(x, \log(y))$ une loi uniforme sur $\mathbb{R}^{p+1}$, toutes les composantes étant indépendantes. Ceci correspond en fait au cas où B^{-1}, υ et ν tendent vers 0 ; dans ce cas, il faut évidemment aussi que $r \to 0$. En dehors de ce cas limite, on a généralement $r = p$ et les densités existent et sont « propres » (elles s'intègrent à 1). ||

La loi marginale multivariée de x est une loi de Student multivariée qui est définie au point suivant. On peut montrer que :

$$(x, y) \sim N_p\gamma(\mu, B, \upsilon, \nu) \Rightarrow x \sim t_p\,(\mu, \upsilon B, \nu)\,. \tag{2A.72}$$

Lorsque $p = 1$, on obtient la densité normale-gamma pour le couple (x, y).

A.4.7 Densité Student multivariée

A.4.7.1 Définition

La densité Student p-variée est définie comme le mélange d'une loi normale p-variée par une loi gamma particulière. C'est en fait la densité marginale de $x \in \mathbb{R}^p$ dans une normale-gamma multivariée particulière où le paramètre υ de la gamma est égal à 1 :

$$p(x) = f_{t_p}(x \mid \mu, V, \nu) = \int_0^{\infty} f_N\left(x \mid \mu, y^{-1}V\right) f_{\gamma}\left(y \middle| \frac{\nu}{2}, \frac{\nu}{2}\right) dy \tag{2A.73}$$

$$= \int_0^{\infty} f_{N\gamma}(x, y \mid \mu, V, v = 1, \nu)\, dy$$

$$= \frac{|V|^{-1/2} (\nu/2)^{\nu/2}}{(2\pi)^{p/2}\Gamma(\nu/2)} \int_0^{\infty} y^{p/2+\nu/2-1}\, e^{-y\left(\nu/2+(1/2)(x-\mu)'V^{-1}(x-\mu)\right)}\, dy$$

$$= \frac{\nu^{\nu/2} |V|^{-1/2}\, \Gamma\left((\nu+p)/2\right)}{\pi^{p/2}\Gamma(\nu/2)} \left[\nu + (x-\mu)'V^{-1}(x-\mu)\right]^{-(\nu+p)/2} \tag{2A.74}$$

écrira plus simplement :

$$x \sim t_p(\mu, V, \nu)\,. \tag{2A.75}$$

On peut montrer que les moments de x sont :

$$\begin{cases} \mathrm{E}(x) = \mu \\ \mathrm{Var}(x) = \dfrac{\nu}{\nu-2} V \quad (\nu > 2)\,. \end{cases} \tag{2A.76}$$

Notons que, dans le cas particulier $p = 1$, on retrouve la densité Student univariée décrite haut en (2A.34), en se rappelant que $\pi^{1/2} = \Gamma(1/2)$.

2A.7.2 Marginales et conditionnelles

A partir des propriétés de la normale p-variée (2A.57), on retrouve immédiatement les propriétés des marginales et conditionnelles d'une Student p-variée. Si on partitionne x de la façon suivante : $x = (x_1', x_2')' \in \mathbb{R}^p$ où $x_i \in \mathbb{R}^{p_i}$, $i = 1, 2$ et $p_1 + p_2 = p$. En partitionnant μ et V comme x :

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}, \tag{2A.77}$$

on obtient :

$$x_1 \sim t_{p_1}(\mu_1, V_{11}, \nu) \tag{2A.78}$$

$$x_2 \mid x_1 \sim t_{p_2}(\mu_{2\cdot1}, V_{22\cdot1}, \nu + p_1) \tag{2A.79}$$

où :

$$\mu_{2\cdot1} = \mu_2 + V_{21}(V_{11})^{-1}(x_1 - \mu_1) \tag{2A.80}$$

$$V_{22\cdot1} = V_{22} - V_{21}(V_{11})^{-1}V_{12}\,. \tag{2A.81}$$

Marginalement, chaque composante x_i, $i = 1, \ldots, p$ de x est donc une Student univariée :

$$x_i \sim t(\mu_i, v_{ii}, \nu)\,, \tag{2A.82}$$

où v_{ii} est l'élément (i, i) de la matrice V.

A.4.7.3 Transformations linéaires

Les propriétés des transformations linéaires de normales multivariées se transmettent, par définition de la Student, sur les variables Student multivariées.

Soit $x \sim t_p(\mu, V, \nu)$ et $y = Ax + b$ où $A : (r \times p)$ et $b = (r \times 1)$ avec $r \leqslant p$. En utilisant (2A.55) et (2A.73), on obtient immédiatement :

$$y \sim t_r \left(A\mu + b, AVA', \nu \right). \tag{2A.83}$$